A Probabilistic Approach to Multivariable Robust Filtering and Open-Loop Control

Kenth Öhrn, Anders Ahlén, *Senior Member, IEEE*, and Mikael Sternad, *Senior Member, IEEE*

***Abstract*—A new approach to robust filtering, prediction, and smoothing of discrete-time signal vectors is presented. Linear time-invariant filters are designed to be insensitive to spectral uncertainty in signal models. The goal is to obtain a simple design method, leading to filters which are not overly conservative. Modeling errors are described by sets of models, parameterized by random variables with known covariances. These covariances could either be estimated from data or be used as robustness "tuning knobs." A robust design is obtained by minimizing the $\mathcal{H}_2$-norm or, equivalently, the mean square estimation error, averaged with respect to the assumed model errors. A polynomial solution, based on an averaged spectral factorization and a unilateral Diophantine equation, is derived. The robust estimator is referred to as a cautious Wiener filter. It turns out to be only slightly more complicated to design than an ordinary Wiener filter. The methodology can be applied to any open-loop filtering or control problem. In particular, we illustrate this for the design of robust multivariable feedforward regulators, decoupling and model matching filters.**

I. Introduction

FOR any model-based filter, modeling errors are a potential source of performance degradation. Here, we will propose a cautious Wiener filter for the prediction, filtering, or smoothing of discrete-time signal vectors. As in the scalar case, discussed in [36], it constitutes a generalization of the polynomial equations methodology pioneered by Kučera [21]. The design is based on a stochastic description of model errors, with relations to e.g., the stochastic embedding concept of Goodwin and coworkers [11], [12]. To be more specific, our problem formulation is as follows:

- A set of (true) dynamic systems is assumed to be well described by a set of discrete-time, stable, linear and time-invariant transfer function matrices

$$\mathcal{F} = \mathcal{F}_o + \Delta\mathcal{F}. \tag{1.1}$$

  We call such a set an extended design model, in which $\mathcal{F}_o$ represents a stable nominal model, while an error model $\Delta\mathcal{F}$ describes a set of stable transfer functions, parameterized by stochastic variables. The random variables enter linearly into $\Delta\mathcal{F}$, and they are assumed independent of the noise.

- A single robust linear filter is to be designed for the whole class of possible systems. Robust performance is obtained by minimizing the averaged mean square estimation error criterion

$$J = \text{trace}\bar{E}E(\varepsilon(k)\varepsilon(k)^*). \tag{1.2}$$

  Here, $\varepsilon(k)$ is the estimation error vector, E denotes expectation over noise and $\bar{E}$ is an expectation over the stochastic variables parameterizing the error model $\Delta\mathcal{F}$.

The averaged mean square error has been used previously in the literature by e.g., Chung and Bélanger [9], Speyer and Gustafson [32], and by Grimble [13]. These works were based on assumptions of small parametric uncertainties and on series expansions of uncertain parameters. We suggest the use of the criterion (1.2), together with a particular description of the set (1.1): transfer function elements in $\Delta\mathcal{F}$ have stochastic numerators and fixed denominators. Such models can describe nonparametric uncertainty and undermodeling as well as parametric uncertainty. A discussion of the utility, and versatility, of linearly parameterized stochastic error models can be found in [36].

Most previous suggestions for obtaining robust filters have been based on some type of minimax approach [10], [24]. A paper [26] by Martin and Mintz takes both spectral uncertainty and uncertainty in the noise distribution into account. The resulting filter will, however, be of very high order. Minimax design of a filter $\mathcal{R}$ becomes very complex, unless there exists either a saddle point or a boundary point solution. A crucial condition here is that $\min_{\mathcal{R}} \max_{\mathcal{F}}$ equals $\max_{\mathcal{F}} \min_{\mathcal{R}}$. If so, one can search for models whose optimal filter gives the worst (nominal) performance and use the corresponding filter. As compared to finding the worst case with respect to a set of models, this is a much simpler task. It can still, however, be computationally demanding. See [19], [28], [31], [38], and the survey paper by Kassam and Poor [20]. The condition $\min_{\mathcal{R}} \max_{\mathcal{F}} = \max_{\mathcal{F}} \min_{\mathcal{R}}$ is not fulfilled in numerous problems, which makes them very difficult to solve. See, e.g., Example 5 in [36] and the example in Section IV.

Kalman filter-like estimators have recently been developed for systems with structured and possibly time-varying parametric uncertainty of the type

$$x(k+1) = (\mathbf{A} + \mathbf{D}\mathbf{\Delta}(k)\mathbf{E})x(k) + w(k)$$

where the matrix $\mathbf{\Delta}(k)$ contains norm-bounded uncertain parameters. See [30], [7], and [39] for continuous-time results and [40] for the discrete-time one-step predictor. See also [16]

Manuscript received July 2, 1993; revised May 20, 1994. Recommended by Past Associate Editor, B. Pasik-Duncan. The work was supported in part by the Swedish Research Council for Engineering Sciences (TFR) under Grant 92-775.

The authors are with the Systems and Control Group, Uppsala University, P.O. Box 27, S-751 03 Uppsala, Sweden.

IEEE Log Number 9408271.

for a related method. For systems which are stable for all $\Delta(k)$, an upper bound on the estimation error covariance matrix can be minimized by solving two coupled Riccati equations, combined with a one-dimensional numerical search. This represents a computational simplification, as compared to previous minimax designs. Still, the resulting estimators are quite conservative, partly because they rest on worst case design. This conservativism is illustrated and discussed in [29] and [37].

The method suggested in the present paper is computationally simpler than any of the minimax schemes referred to above. It also avoids two drawbacks of worst case designs. First, the stochastic variables in $\Delta\mathcal{F}$ need not have compact support. Thus, the descriptions of model uncertainties may have "soft" bounds. These are more readily obtainable in a noisy environment than the hard bounds required for minimax design. Second, not only the range of the uncertainties, but also their likelihood is taken into account by using the expectation $\bar{E}(\cdot)$ of the MSE. Highly probable model errors will affect the estimator design more than do very rare "worst cases." Therefore, the performance loss in the nominal case, the price paid for robustness, becomes smaller than for a minimax design. In other words, conservativeness is reduced. There do exist applications where a worst-case design is mandatory, e.g., for safety reasons. We believe, however, that the average performance of estimators is often a more appropriate measure of performance robustness.

In the present paper, one of our goals will be to present transparent design equations and to hold their number to a minimum without sacrificing numerical accuracy. We use matrix fraction descriptions with diagonal denominators and common denominator forms. This leads to a solution which is, in fact, significantly simpler and numerically better behaved than the corresponding nominal $\mathcal{H}_2$-designs (without uncertainty) presented in [1] or [14]. Somewhat surprisingly, taking model uncertainty into account does not require any new types of design equations. We end up with just two equations for robust estimator design: a polynomial matrix spectral factorization and a unilateral Diophantine equation. The solution provides structural insight; important properties of a robust estimator are evident by direct inspection of the filter expression.

This paper is organized as follows. The filtering problem, model structure (1.1), and criterion (1.2) are discussed in more detail in Section II. Section III presents the design equations and some tools for performance evaluation. The design procedure is illustrated by a thorough numerical example in Section IV. The resulting estimator reduces the impact of model uncertainty and limited signal energy by using multiple sensors in an efficient way. In Section V the design of robust feedforward regulators, servos, and model matching filters is discussed.

Remarks on the Notation: Signals and polynomial coefficients may, in the following, be complex valued. (This is required in, e.g., communications applications.) Let p_j^* denote the complex conjugate (and transpose for matrices) of a polynomial coefficient p_j. For any polynomial

$$P(q^{-1}) = p_o + p_1 q^{-1} + \cdots + p_{np} q^{-np}$$

in the backward shift operator q^{-1}, define the conjugate polynomial

$$P_*(q) \triangleq p_o^* + p_1^* q + \cdots + p_{np}^* q^{np}$$

where q is the forward shift operator. A polynomial $P(q, q^{-1})$ having coefficients of both q and q^{-1} will be called double sided. Rational matrices, or transfer functions, are denoted by boldface calligraphic symbols, e.g., $\boldsymbol{\mathcal{R}}(q^{-1})$. Polynomial matrices are denoted by boldface symbols, such as $\boldsymbol{P}(q^{-1})$, while constant matrices are denoted as $\mathbf{P}$. For example, the identity matrix of dimension n is denoted $\mathbf{I}_n$. We denote the trace of $\boldsymbol{P}$ by $\mathrm{tr}\boldsymbol{P}$. For polynomial or rational matrices, $\boldsymbol{P}_*(q)$ and $\boldsymbol{\mathcal{R}}_*(q)$ means complex conjugate, transpose, and substitution of q for q^{-1}. When appropriate, the complex variable z or $e^{i\omega}$ is substituted for the forward shift operator q. Arguments of polynomials and matrices are often omitted, when there is no risk of misunderstanding. The degree of a polynomial matrix is the highest degree of any of its polynomial elements. Square polynomial matrices $\boldsymbol{P}(q^{-1})$ are called stable if all zeros of $\det \boldsymbol{P}(z^{-1})$ are located in $|z| < 1$. A rational matrix is defined as stable if all its elements are stable. Causality is defined in the same way.

A rational matrix $\boldsymbol{\mathcal{G}}(q^{-1})$ may be represented by polynomial matrices as a matrix fraction description (MFD), either left $\boldsymbol{\mathcal{G}} = \boldsymbol{A}_1^{-1}\boldsymbol{B}_1$ or right $\boldsymbol{\mathcal{G}} = \boldsymbol{B}_2\boldsymbol{A}_2^{-1}$. It may also be represented in a common denominator form $\boldsymbol{\mathcal{G}} = \boldsymbol{B}/A$, where $\boldsymbol{B}$ is a polynomial matrix. The scalar and monic polynomial A is then the least common denominator of all elements in $\boldsymbol{\mathcal{G}}$. Denominator matrices in MFD's are assumed to have identity matrices as leading coefficients of their matrix polynomial representations, thus $\boldsymbol{A}_i(0) = \mathbf{I}$ above.

II. The Robust Estimation Problem

Consider the following extended design model

$$\begin{aligned} y(k) &= \boldsymbol{\mathcal{G}}(q^{-1})u(k) + \boldsymbol{\mathcal{H}}(q^{-1})v(k) \\ u(k) &= \boldsymbol{\mathcal{F}}(q^{-1})e(k) \\ f(k) &= \boldsymbol{\mathcal{D}}(q^{-1})u(k) \end{aligned} \tag{2.1}$$

where $\boldsymbol{\mathcal{G}}$, $\boldsymbol{\mathcal{H}}$, $\boldsymbol{\mathcal{F}}$, and $\boldsymbol{\mathcal{D}}$ are stable and causal, but possibly uncertain, transfer functions of dimension $p|s$, $p|r$, $s|n$, and $\ell|s$, respectively. The noise sequences $\{e(k)\}$ and $\{v(k)\}$ are mutually uncorrelated and zero mean stochastic sequences. To obtain a simple notation they are assumed to have unit covariance matrices, so scaling and uncertainty of the covariances are included in $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{H}}$, respectively. The signal $y(k)$ is assumed measurable, while $f(k)$ is the signal to be estimated.

A. Multisignal Estimation

From data $y(k)$ up to time $k+m$, an estimator

$$\hat{f}(k|k+m) = \boldsymbol{\mathcal{R}}(q^{-1})y(k+m) \tag{2.2}$$

of $f(k)$ is sought. See Fig. 1. The estimator may be a predictor ($m < 0$), a filter ($m = 0$), or a fixed lag smoother ($m > 0$). Here $\boldsymbol{\mathcal{R}}$, of dimension $\ell|p$, is required to be stable and causal.

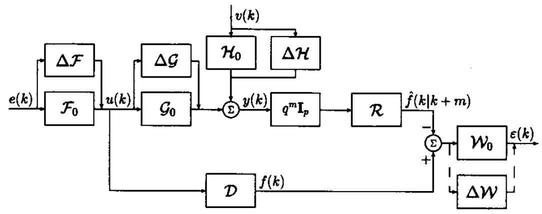


Fig. 1. A general linear filtering problem formulation. Based on noisy measurements $y(k+m)$, the signal $f(k)$is to be estimated. Model errors in transfer functions are described by stochastic error models.

The transfer function $\mathcal{R}$ is designed to minimize the averaged mean square error (MSE) criterion (1.2)

$$J = \operatorname{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^*) = \bar{E}E(\varepsilon(k)^*\varepsilon(k)) = \sum_{i=1}^{\ell} \bar{E}E|\varepsilon_i(k)|^2 \tag{2.3}$$

where

$$\varepsilon(k) = (\varepsilon_1(k)\cdots\varepsilon_\ell(k))^T \triangleq \mathcal{W}(q^{-1})(f(k) - \hat{f}(k|k+m)).$$

Above, $\mathcal{W}$ is a stable and causal $\ell|\ell$ rational weighting matrix, with a stable and causal inverse. It may be used by the designer to emphasize filtering performance in particular frequency bands. In filtering problems, $\mathcal{W}$ is not assumed uncertain.

Model (2.1) offers considerable flexibility. For example, when estimating a signal $u(k)$ in colored noise, we set $\mathcal{G} = \mathcal{D} = \mathbf{I}_s$, giving $f(k) = u(k)$. In deconvolution, or input estimation problems, $\mathcal{G}$ is a dynamic system and $\mathcal{D} = \mathbf{I}_s$. In a state estimation problem, $u(k)$ is the state vector, $\mathcal{G}$ and $\mathcal{D}$ are constant matrices while $\mathcal{H}v(k)$ represents (colored) measurement noise. Other special cases are discussed in [2], [3], and [8].

Example: An application where uncertain dynamics in $\mathcal{G}$ is of interest is equalizer design for digital mobile radio communications [23]. A signal $u(k)$ then propagates along multiple paths, with different time delays, represented by delays in $\mathcal{G}$. The receiving antenna may have $p > 1$ elements (diversity design). See, e.g., [4]. Thus, an appropriate model of $\mathcal{G}$ is a column vector of FIR channels, i.e., a vector of polynomials. The polynomials coefficients are estimated from short and noisy training sequences, with a known input $\{u(k)\}$. Estimation errors are inevitable. The task of a (robust) equalizer is to estimate $u(k)$, based on noisy measurements $y(k+m)$, a nominal model $\mathcal{G}_o$, and an estimate of the amount of model uncertainty □

B. Parameterization of the Model

We choose to parameterize $\mathcal{G}$ and $\mathcal{H}$ as left MFD's having diagonal denominators,[1] while $\mathcal{F}, \mathcal{D}$, and $\mathcal{W}$ are parameterized in common denominator form

$$\mathcal{G} = \boldsymbol{A}^{-1}\boldsymbol{B}; \quad \mathcal{H} = \boldsymbol{N}^{-1}\boldsymbol{M} \tag{2.4}$$

$$\mathcal{F} = \frac{1}{D}\boldsymbol{C}; \quad \mathcal{D} = \frac{1}{T}\boldsymbol{S}; \quad \mathcal{W} = \frac{1}{U}\boldsymbol{V}.$$

[1] Note that this is a natural choice, if transfer functions are obtained by means of identification.

We have made these choices to obtain tidy and transparent design equations and to avoid coprime factorizations, which are known to be numerically sensitive. In (2.4), it is assumed that $\mathcal{G}, \mathcal{H}$, and $\mathcal{F}$ may be uncertain. Introduction of uncertainty in the weighting matrix $\mathcal{W}$ is not motivated in filtering problems. Its role in open-loop control will be discussed in Section III-B below. It is shown in Appendix C that uncertainty in $\mathcal{D}$ does not affect the optimal filter design, provided it is uncorrelated to uncertainties in other blocks. Therefore, uncertainty in $\mathcal{D}$ is not introduced.

The extended design models, cf. (1.1) and (2.1)

$$\mathcal{G} = \mathcal{G}_o + \Delta\mathcal{G}, \quad \mathcal{H} = \mathcal{H}_o + \Delta\mathcal{H}, \quad \mathcal{F} = \mathcal{F}_o + \Delta\mathcal{F}$$

are now expressed in polynomial matrix form. Using $\hat{\boldsymbol{B}}_o = \boldsymbol{A}_1\boldsymbol{B}_o$, $\hat{\boldsymbol{B}}_1 = \boldsymbol{A}_o\boldsymbol{B}_1$ etc. we introduce

$$\begin{aligned}
\mathcal{G} &= \boldsymbol{A}_o^{-1}\boldsymbol{B}_o + \boldsymbol{A}_1^{-1}\boldsymbol{B}_1\Delta\boldsymbol{B} \\
&= \boldsymbol{A}_o^{-1}\boldsymbol{A}_1^{-1}(\hat{\boldsymbol{B}}_o + \hat{\boldsymbol{B}}_1\Delta\boldsymbol{B}) \triangleq \boldsymbol{A}^{-1}\boldsymbol{B} \\
\mathcal{H} &= \boldsymbol{N}_o^{-1}\boldsymbol{M}_o + \boldsymbol{N}_1^{-1}\boldsymbol{M}_1\Delta\boldsymbol{M} \\
&= \boldsymbol{N}_o^{-1}\boldsymbol{N}_1^{-1}(\hat{\boldsymbol{M}}_o + \hat{\boldsymbol{M}}_1\Delta\boldsymbol{M} \triangleq \boldsymbol{N}^{-1}\boldsymbol{M} \\
\mathcal{F} &= \frac{1}{D_o}\boldsymbol{C}_o + \frac{1}{D_1}\boldsymbol{C}_1\Delta\boldsymbol{C} \\
&= \frac{1}{D_oD_1}(\hat{\boldsymbol{C}}_o + \hat{\boldsymbol{C}}_1\Delta\boldsymbol{C}) \triangleq \frac{1}{D}\boldsymbol{C}.
\end{aligned} \tag{2.5}$$

Above, $\mathcal{G}_o = \boldsymbol{A}_o^{-1}\boldsymbol{B}_o$ represents the nominal model and $\Delta\mathcal{G} = \boldsymbol{A}_1^{-1}\boldsymbol{B}_1\Delta\boldsymbol{B}$ the error model. The same holds for $\mathcal{H}$ and $\mathcal{F}$. The diagonal polynomial matrices $\boldsymbol{A} = \boldsymbol{A}_o\boldsymbol{A}_1$, $\boldsymbol{N} = \boldsymbol{N}_o\boldsymbol{N}_1$, and the polynomials $D = D_oD_1, T$ and U are all assumed to be stable, with causal inverses. Denominator polynomials are assumed monic. In the error models, the polynomial D_1, the diagonal matrices $\boldsymbol{A}_1$ and $\boldsymbol{N}_1$, and the matrices $\boldsymbol{C}_1$, $\boldsymbol{B}_1$ and $\boldsymbol{M}_1$ are fixed. They can be used to tailor the error models for specific needs. For example, if multiplicative error models are deemed appropriate, we use $\boldsymbol{A}_1 = \boldsymbol{A}_o$, $\boldsymbol{B}_1 = \boldsymbol{B}_o\boldsymbol{B}_m$ etc., with $\boldsymbol{B}_m$ to be specified.

The matrices $\Delta\boldsymbol{B}$, $\Delta\boldsymbol{C}$, and $\Delta\boldsymbol{M}$ contain polynomials, with jointly distributed random variables as coefficients. These coefficients parameterize the class of assumed true systems. One particular modeling error is represented by one particular realization of the random coefficients.[2] Element ij of a stochastic polynomial matrix $\Delta\boldsymbol{P}$ is denoted

$$\Delta P^{ij} \triangleq [\Delta\boldsymbol{P}]_{ij} = \Delta p_o^{ij} + \Delta p_1^{ij}q^{-1} + \cdots + \Delta p_{\delta p}^{ij}q^{-\delta p} \tag{2.6}$$

where δp is the degree of $\Delta\boldsymbol{P}$, i.e., the highest degree appearing in any polynomial ΔP^{ij}. All coefficients have zero means, so the nominal model is the average model in the set. Only the second-order moments of the random coefficients need to be specified, since the type of distribution, and higher order moments, will not affect the filter design. The parameter

[2] For a given system realization, the coefficients are assumed time-invariant and independent of the time-series $e(k)$ and $v(k)$. This is in contrast to the approach of Haddad and Bernstein in [15], who represent the effect of uncertainties by multiplicative noises. For a given uncertainty variance, a noise representation would underestimate the true effect of (time-invariant) parameter deviations in the dynamics.

covariances are denoted $\bar{E}(\Delta p_r^{ij})(\Delta p_s^{\ell k})^*$ and are collected in covariance matrices $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$; see Section II-C.

We now introduce the assumption

- **A1.** The coefficients of all polynomial elements in ΔC are independent of those in ΔB.

It is possible to exclude Assumption A1, but it does simplify the solution, and it is also reasonable in most practical cases.

Error models can be obtained from ordinary identification experiments, provided the model structures match. For SISO systems, error models can be estimated in presence of under-modeling, using a maximum likelihood approach [11]. Even if the statistics is hard to obtain, one could still use the elements of covariance matrices pragmatically, as robustness "tuning knobs." They are then used similarly as when weighting matrices are adjusted in LQG controller design. An objective could be to obtain reasonable performance for the uncertainty set, for a prespecified acceptable degradation of performance in the nominal case. The error models may also be used to account for a slowly time-varying dynamics [25].

One way of obtaining the models (2.5)–(2.6) is by series expansion of state-space models with parametric uncertainty [37]. Parameter deviations are represented by stochastic variables. For small uncertainties, a first-order expansion can be used, which will directly lead to models of type (2.5). For larger uncertainties, a second-order Taylor expansion is usually sufficient; see [29]. Error models for nonparametric uncertainties can be adjusted directly to frequency domain data. In that context, a very useful concept is provided by the stochastic frequency domain theory of Goodwin and Salgado; see [12].

C. Covariance Matrices for the Stochastic Coefficients

To represent the uncertainties of the system in a natural way, covariance matrices will be organized as follows. The ijth element of a stochastic polynomial matrix ΔP can be expressed as

$$\Delta P^{ij}(q^{-1}) = \varphi^T(q^{-1})\bar{p}_{ij} \tag{2.7}$$

where

$$\varphi^T(q^{-1}) = (1\, q^{-1}\cdots q^{-\delta p})\ ;\ \bar{p}_{ij} = (\Delta p_o^{ij}\Delta p_1^{ij}\cdots\Delta p_{\delta p}^{ij})^T. \tag{2.8}$$

The cross covariance matrix $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$, of dimension $\delta p+1|\delta p+1$, between coefficients of $\Delta P^{ij}(q^{-1})$ and $\Delta P^{\ell k}(q^{-1})$, is given by

$$\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)} = \bar{E}\,\bar{p}_{ij}\bar{p}_{\ell k}^* = \begin{bmatrix} \bar{E}(\Delta p_o^{ij})(\Delta p_o^{\ell k})^* & \cdots & \bar{E}(\Delta p_o^{ij})(\Delta p_{\delta p}^{\ell k})^* \\ \vdots & \ddots & \vdots \\ \bar{E}(\Delta p_{\delta p}^{ij})(\Delta p_o^{\ell k})^* & \cdots & \bar{E}(\Delta p_{\delta p}^{ij})(\Delta p_{\delta p}^{\ell k})^* \end{bmatrix} \tag{2.9}$$

where $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,ij)}$ is Hermitian and positive semidefinite, while $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)} = (\mathbf{P}_{\Delta\mathbf{P}}^{(\ell k,ij)})^*$. Thus

$$\bar{E}(\Delta P^{ij}\Delta P_*^{\ell k}) = \bar{E}(\varphi^T(q^{-1})\bar{p}_{ij}\bar{p}_{\ell k}^*\varphi_*^T(q)) = \varphi^T\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}\varphi_*^T. \tag{2.10}$$

With autocovariances, $(ij) = (\ell k)$, we model the uncertainty within each input–output pair. Cross-dependencies between different transfer functions may also be known. For example, uncertainty in one single physical parameter may very well enter into several transfer functions between inputs and outputs. Such effects are captured by cross covariances, $(ij) \neq (\ell k)$.

We collect all matrices of type (2.9) into one large covariance matrix, organized as shown by (2.11) at the bottom of the page. If ΔP has dimension $n|m$, then $\mathbf{P}_{\Delta\mathbf{P}}$ is composed of nm by nm covariance matrices $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$. The structure of (2.11) is useful from a design point of view. If, for example, a multivariable moving average model, or FIR model, is to be identified, then (2.11) is the natural way of representing the covariance matrix. If we instead prefer to use the blocks $\mathbf{P}_{\Delta\mathbf{P}}^{(ij,\ell k)}$ of (2.11) as multivariable "tuning knobs," a given amount of uncertainty can be assigned to a specific input–output pair.

III. Design Of Robust Filters

A. An Averaged Spectral Factorization

We define an averaged spectral factor $\beta(q^{-1})$ as the numerator polynomial matrix of an averaged innovations model. It constitutes a key element of the robust filter. The average, over the set of models, of the spectral density matrix $\Phi_y(e^{i\omega})$ of the measurement $y(k)$ is given by

$$\bar{E}\{\Phi_y(e^{i\omega})\} = \frac{1}{DD_*}A^{-1}N^{-1}\beta\beta_*N_*^{-1}A_*^{-1}.$$

$$\mathbf{P}_{\Delta\mathbf{P}} = \begin{pmatrix} \begin{bmatrix} \mathbf{P}_{\Delta\mathbf{P}}^{(11,11)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(11,1m)} \\ \vdots & \ddots & \vdots \\ \mathbf{P}_{\Delta\mathbf{P}}^{(1m,11)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(1m,1m)} \end{bmatrix} & \cdots & \begin{bmatrix} \mathbf{P}_{\Delta\mathbf{P}}^{(11,n1)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(11,nm)} \\ \vdots & & \vdots \\ \mathbf{P}_{\Delta\mathbf{P}}^{(1m,n1)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(1m,nm)} \end{bmatrix} \\ \vdots & \ddots & \vdots \\ \begin{bmatrix} \mathbf{P}_{\Delta\mathbf{P}}^{(n1,11)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(n1,1m)} \\ \vdots & & \vdots \\ \mathbf{P}_{\Delta\mathbf{P}}^{(nm,11)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(nm,1m)} \end{bmatrix} & \cdots & \begin{bmatrix} \mathbf{P}_{\Delta\mathbf{P}}^{(n1,n1)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(n1,nm)} \\ \vdots & \ddots & \vdots \\ \mathbf{P}_{\Delta\mathbf{P}}^{(nm,n1)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(nm,nm)} \end{bmatrix} \end{pmatrix} \tag{2.11}$$

The square polynomial matrix $\beta(z^{-1})$ is given by the stable solution to

$$\beta\beta_* = \bar{E}\{NBCC_*B_*N_* + DAMM_*A_*D_*\}. \quad (3.1)$$

Note that N^{-1} and A^{-1} are diagonal and will thus commute. The averaged second-order statistics of $y(k)$ is thus described by the same spectral density as for a vector-ARMA model

$$\bar{y}(k) = \frac{1}{D}A^{-1}N^{-1}\beta\,\epsilon(k) \quad (3.2)$$

where $\epsilon(k)$ is white with a unit covariance matrix. This model is denoted as the averaged innovations model. (Note that $\bar{y}(k) \neq y(k)$, but $\Phi_{\bar{y}}(e^{i\omega}) = \bar{E}\{\Phi_y(e^{i\omega})\}$). When constructing the right-hand side of (3.1), the following results are useful.

Lemma 1: Let $H(q,q^{-1})$ be an $m|m$ polynomial matrix with double-sided polynomial elements having stochastic coefficients. Also, let $G(q^{-1})$ be an $n|m$ polynomial matrix with polynomial elements having stochastic coefficients, independent of all those of H. Then

$$\bar{E}[GHG_*] = \bar{E}[G\bar{E}(H)G_*]. \quad (3.3)$$

□

Proof: See Appendix A.

Now, introduce the double-sided polynomial matrices

$$\tilde{C}\tilde{C}_* \triangleq \bar{E}(CC_*); \quad \tilde{B}_C\tilde{B}_{C*} \triangleq \bar{E}(B\tilde{C}\tilde{C}_*B_*);$$
$$\tilde{M}\tilde{M}_* \triangleq \bar{E}(MM_*). \quad (3.4)$$

Invoking (2.5) and using the fact that the stochastic coefficients are assumed to be zero mean, gives

$$\tilde{C}\tilde{C}_* = \hat{C}_o\hat{C}_{o*} + \hat{C}_1\bar{E}(\Delta C\Delta C_*)\hat{C}_{1*}$$
$$\tilde{B}_C\tilde{B}_{C*} = \hat{B}_o\tilde{C}\tilde{C}_*\hat{B}_{o*} + \hat{B}_1\bar{E}(\Delta B\tilde{C}\tilde{C}_*\Delta B_*)\hat{B}_{1*}$$
$$\tilde{M}\tilde{M}_* = \hat{M}_o\hat{M}_{o*} + \hat{M}_1\bar{E}(\Delta M\Delta M_*)\hat{M}_{1*}. \quad (3.5)$$

Factorizations to obtain $\tilde{C}$, $\tilde{B}_C$, etc. need not be performed. The double-sided polynomial matrices are expressed as $\tilde{C}\tilde{C}_*$, etc. merely to simplify the notation.

Lemma 2: Let Assumption A1 hold. By using (3.4), (3.5) and invoking Lemma 1, the averaged spectral factorization (3.1) can be expressed as

$$\beta\beta_* = N\tilde{B}_C\tilde{B}_{C*}N_* + DA\tilde{M}\tilde{M}_*A_*D_*. \quad (3.6)$$

□

Proof: See Appendix A.

With a given right-hand side, (3.6) is just an ordinary polynomial matrix left spectral factorization. It is solvable under the following mild assumption

- **A2.** The averaged spectral density matrix $\bar{E}\{\Phi_y(e^{i\omega})\}$ is nonsingular for all ω.

This assumption is equivalent to the right-hand side of (3.6) being nonsingular on $|z| = 1$. Then, the solution to (3.6) is unique, up to a right unitary factor. (If $HH_* = \mathbf{I}$, then $\beta\beta_* = (\beta H)(H_*\beta_*)$.) Under Assumption A2, a solution exists, with β having nonsingular leading coefficient matrix $\beta(0)$. Its degree, $n\beta$, will be determined by the maximal degree of the two right-hand side terms of (3.6).[3]

To obtain the right-hand side of (3.6), averaged polynomial matrices $\bar{E}(\Delta PH\Delta P_*)$ have to be computed, where $H(q,q^{-1}) = \tilde{C}\tilde{C}_*$ or $\mathbf{I}$. It is shown in Appendix B that the ijth element of $\bar{E}(\Delta PH\Delta P_*)$ is given by

$$\bar{E}[\Delta PH\Delta P_*]_{ij} = \mathrm{tr}H\begin{bmatrix}\varphi^T & & 0\\ & \ddots & \\ 0 & & \varphi^T\end{bmatrix} \times \begin{bmatrix}\mathbf{P}_{\Delta\mathbf{P}}^{(i1,j1)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(im,j1)}\\ \vdots & \ddots & \vdots\\ \mathbf{P}_{\Delta\mathbf{P}}^{(i1,jm)} & \cdots & \mathbf{P}_{\Delta\mathbf{P}}^{(im,jm)}\end{bmatrix}\begin{bmatrix}\varphi_*^T & & 0\\ & \ddots & \\ 0 & & \varphi_*^T\end{bmatrix} \quad (3.7)$$

where φ^T was defined in (2.8). The block covariance matrix in (3.7) constitutes the block-transpose of the ijth block $[\,\cdot\,]$ of $\mathbf{P}_{\Delta\mathbf{P}}$ in (2.11). Average factors in (3.5) are readily obtained by substituting ΔC, ΔB, and ΔM for ΔP in (3.7).

B. A Second Spectral Factorization

In the feedforward control problems discussed in Section V, we shall allow for W being uncertain. Using a common denominator form, W is parameterized in a similar way as $\mathcal{F}$ (cf. (2.5))

$$W = V_o\frac{1}{U_o} + \Delta VV_1\frac{1}{U_1} = \frac{1}{U_oU_1}(\hat{V}_o + \Delta V\hat{V}_1) \triangleq \frac{1}{U}V. \quad (3.8)$$

A stable square matrix $\tilde{V}$, with $\tilde{V}(0)$ nonsingular, is introduced as a solution of the right spectral factorization

$$\tilde{V}_*\tilde{V} = \bar{E}(V_*V) = \hat{V}_{o*}\hat{V}_o + \hat{V}_{1*}\bar{E}(\Delta V_*\Delta V)\hat{V}_1. \quad (3.9)$$

Also, introduce the following assumptions

- **A3.** The coefficients of ΔV are independent of all other stochastic coefficients.
- **A4.** The right-hand side of (3.9) is nonsingular on the unit circle.

Whenever W is known, ($W = V_o/U_o = V/U$), (3.9) need not be solved, and $\tilde{V} = V_o = V$. This will be the case in filtering problems.

C. The Cautious Multivariable Wiener filter

Theorem 1: Assume an extended design model (2.1), (2.4), (2.5), (3.8) to be given, with known covariance matrices (2.11). Assume A1–A4 to hold. A realizable estimator of $f(k)$ then minimizes the averaged MSE (2.3), among all linear time-invariant estimators based on $y(k+m)$, if and only if it has the same coprime factors as

$$\hat{f}(k|k+m) = \mathcal{R}\,y(k+m) = \frac{1}{T}\tilde{V}^{-1}Q\beta^{-1}NA\,y(k+m). \quad (3.10)$$

[3] When solving (3.6), we have utilized an algorithm by Ježek and Kučera, presented in [18]. It provides a solution with an upper triangular full rank leading coefficient matrix. For an overview of spectral factorization algorithms, see [22].

Here, $\boldsymbol{\beta}(q^{-1})$ is obtained from (3.6), $\tilde{\boldsymbol{V}}(q^{-1})$ from (3.9) while $\boldsymbol{Q}(q^{-1})$ together with $\boldsymbol{L}_*(q)$, both of dimensions $\ell|p$, is the unique solution to the unilateral Diophantine equation

$$q^{-m}\tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_* = \boldsymbol{Q}\boldsymbol{\beta}_* + q\boldsymbol{L}_*UTD\,\mathbf{I}_p \qquad (3.11)$$

with generic[4] degrees

$$\begin{aligned} nQ &= \max(n\tilde{v} + ns + n\tilde{c} + m, nu + nt + nd - 1) \\ nL_* &= \max(n\tilde{c} + n\hat{b}_o + nn - m, n\beta) - 1 \end{aligned} \qquad (3.12)$$

where $ns = \deg \boldsymbol{S}$ etc. When applying the estimator (3.10) on an ensemble of systems, the minimal criterion value becomes

$$\begin{aligned} \mathrm{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^*)_{\min} = \mathrm{tr}\frac{1}{2\pi j}\oint_{|z|=1}\Big\{&\boldsymbol{L}_*\boldsymbol{\beta}_*^{-1}\boldsymbol{\beta}^{-1}\boldsymbol{L} + \frac{1}{UTDD_*T_*U_*} \\ &\times\tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\Big[\mathbf{I}_n - \tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1}\boldsymbol{\beta}^{-1}\boldsymbol{N}\hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\Big] \\ &\times \tilde{\boldsymbol{C}}_*\boldsymbol{S}_*\tilde{\boldsymbol{V}}_*\Big\}\frac{dz}{z}. \end{aligned} \qquad (3.13)$$

□

Proof: See Appendix C.

Remarks: The only new type of computation, as compared to the nominal case described in [1]–[3], is the calculation of averaged polynomials using (3.7).

Since both $\tilde{\boldsymbol{V}}$ and $\boldsymbol{\beta}$ are stable, the estimator $\boldsymbol{\mathcal{R}}$ will be stable.[5] If Assumptions A2 and A4 hold, $\tilde{\boldsymbol{V}}(0)$ and $\boldsymbol{\beta}(0)$ are nonsingular, so $\boldsymbol{\mathcal{R}}$ will be causal.

Note that the diagonal matrix $\boldsymbol{NA} = \boldsymbol{N}_o\boldsymbol{N}_1\boldsymbol{A}_o\boldsymbol{A}_1$ appears explicitly in the filter (3.10). Thus, important properties of the robust estimator are evident by direct inspection. For example, assume some diagonal elements of $\boldsymbol{N}_1^{-1}$ or $\boldsymbol{A}_1^{-1}$ in the error models to have resonance peaks, indicating large uncertainty at the corresponding frequencies. Then, the filter will have notches, so the filter gain from the uncertain components of $y(k+m)$ will be low at the relevant frequencies.

The nominal Wiener filter has as a component a whitening filter. The robust estimator has a similar structure. By multiplying $\boldsymbol{\mathcal{R}}$ by the stable common factor D/D, the filter in (3.10) will contain $\boldsymbol{\beta}^{-1}\boldsymbol{NA}D$ as right factor. This averaged counterpart of a whitening filter is the inverse of the averaged innovations model (3.2).

The model structure (2.4)–(2.5) was selected to obtain a few simple design equations. Other choices are possible, but lead to various complications. For example, if stochastic polynomials had been introduced in the denominators, no exact analytical solution could have been obtained. Also stability would have been a problem. The use of general left MFD representations, instead of forms with diagonal denominators or common denominators, would have led to a solution involving seven coprime factorizations. Such a solution is presented in [29], but it provides less physical insight. It does also exhibit worse numerical behavior, since algorithms for coprime factorization are numerically sensitive.

[4] In special cases, the degrees may be lower.

[5] Stable common factors may exist in (3.10). They could be detected by calculating invariant polynomials of the involved matrices. If such factors have zeros close to the unit circle, it is advisable to cancel them before the filter is implemented. Otherwise, slowly decaying (initial) transients may deteriorate the filtering performance.

Furthermore, (3.11) is a unilateral Diophantine equation, since $\boldsymbol{Q}$ and $\boldsymbol{L}_*$ appear on the same side of the terms in which they are involved. (When the unknowns appear on opposite sides, the equation is bilateral.) This property is a consequence of our choice of U, T, and D as scalar polynomials. Unilateral equations can easily be transformed into a system of linear equations, $\mathbf{AX} = \mathbf{B}$, where $\mathbf{A}$ is a block-Toeplitz matrix. For an example, see Section IV.

Robust design also makes the solution less numerically sensitive. Almost common factors of $\det\boldsymbol{\beta}_*$ and UTD with zeros close to $|z| = 1$ would make the solution of (3.11) numerically sensitive. In the presence of model uncertainty, the risk for this is less than in the nominal case, due to the presence of averaged factors in (3.5). The averaged spectral factor $\boldsymbol{\beta}$ will, in general, have its zeros more distant from the unit circle than the nominal spectral factor, given by (3.20) below. This reduces the numerical difficulty of solving both (3.6) and (3.11).

For every cautious Wiener filter, there exists a system (without uncertainty) for which this estimator is the optimal Wiener filter (see [29]). It is therefore possible to represent model uncertainties by colored noises and then design a Wiener filter for the corresponding system. This correspondence provides a way of understanding the structure of the above desing equations. We do not recommend such an equivalent noise-approach in the actual design, however, for two reasons:

- It is far from trivial to obtain a noise spectrum having similar effect on the filter design as do uncertainties in the block $\boldsymbol{\mathcal{G}}$ in Fig. 1. This is true in particular if the block $\boldsymbol{\mathcal{F}}$ is also uncertain, and if the problem is multivariable.
- It is advantageous from a design point of view to have separate tools which handle different aspects. Error models should represent the effect of modeling uncertainty; noise models should represent disturbances; criterion weighting functions should reflect the priorities of the user. A method which does not distinguish between these aspects will tend to confuse the designer.

The attainable performance improves monotonically with an increasing smoothing lag m. The following result gives the lower bound of the averaged estimation error. This bound can be approached pointwise in the frequency domain for $m < \infty$, by using a criterion filter $\boldsymbol{\mathcal{W}}$ with a high resonance peak.

Corollary 1: The limiting estimator for $m \to \infty$, the nonrealizable cautious Wiener filter, can be expressed as

$$\lim_{m\to\infty} q^m\boldsymbol{\mathcal{R}} = \frac{1}{T}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1}\boldsymbol{\beta}^{-1}\boldsymbol{NA}. \qquad (3.14)$$

Its average performance is given by (3.13) with $\boldsymbol{L} = \mathbf{0}$. If $\boldsymbol{\mathcal{W}} = \mathbf{I}_\ell$, the spectral distribution of the lower bound of the estimation error $f(k) - \hat{f}(k|k+m)$ is

$$\lim_{m\to\infty}\mathrm{tr}\bar{E}\,\boldsymbol{\Phi}_{f-\hat{f}}(e^{i\omega}) = \frac{1}{TDD_*T_*}\mathrm{tr}\{\boldsymbol{S}\tilde{\boldsymbol{C}}\Big[\mathbf{I}_n - \tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1}\boldsymbol{\beta}^{-1}\boldsymbol{N}\hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\Big]\tilde{\boldsymbol{C}}_*\boldsymbol{S}_*\}. \qquad (3.15)$$

The bound can be attained at a frequency ω_1 by an estimator with finite smoothing lag, if it is designed using a weighted criterion where

$$U(e^{-i\omega_1}) \approx 0. \tag{3.16}$$

□

Proof: In a similar way as in Appendix A.3 of [8], it is straightforward to show that $\boldsymbol{L} \to \boldsymbol{0}$ as $m \to \infty$ in (3.11). Thus, (3.11) gives

$$\lim_{m\to\infty} q^m \boldsymbol{Q} = \tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1}. \tag{3.17}$$

The substitution of this expression into (3.10) gives (3.14). The use of $\boldsymbol{L} = \boldsymbol{0}$, $\tilde{\boldsymbol{V}} = \mathbf{I}_\ell$ and $U = 1$ in the integrand of (3.13) gives (3.15). When $U(e^{-i\omega_1}) \approx 0$, we obtain the same effect on the Diophantine equation (3.11) at the frequency ω_1 as if $\boldsymbol{L} \to \boldsymbol{0}$: the rightmost term vanishes. Thus, at ω_1, the gain and the phase of the elements of the polynomial matrix $q^m\boldsymbol{Q}$ are approximately equal to those of (3.17) and the estimation error approaches the lower bound (3.15) ■

Remarks: Note that for realizable estimators (m finite), the lower bound (3.15) is only attainable at distinct frequencies ω_i by means of frequency weighting. For frequencies outside the bandwidth of $\mathcal{W}$, the estimate may be severely degraded. The results of Corollary 1 are illustrated at the end of the example in Section IV.

D. Analytical Expressions for Performance Evaluation

Theorem 2: Let a nominal estimator $\mathcal{R}_n$ be designed based on a nominal model, with no uncertainties. Applying it, instead of (3.10), on an ensemble of systems results in an increase, as compared to (3.13), of the averaged MSE. The increase is given by

$$\begin{aligned}\mathrm{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^*)_n - \mathrm{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^*)_{\min} &= \\ \|\mathcal{W}(\mathcal{R}_n - \mathcal{R})z^m D^{-1}\boldsymbol{A}^{-1}\boldsymbol{N}^{-1}\boldsymbol{\beta}\|_2^2\end{aligned} \tag{3.18}$$

where $\boldsymbol{\beta}$ is defined by (3.1)–(3.6) and $\mathcal{R}$ is the robust estimator (3.10). □

Proof: To obtain (3.18), the nominal filter $\mathcal{R}_n$ is expressed as $\mathcal{R}+(\mathcal{R}_n-\mathcal{R})$. The optimality of $\mathcal{R}$ implies that any modification gives an orthogonal contribution to the criterion. This, and the use of the averaged innovations model (3.2), gives (3.18). Mixed terms vanish, due to the orthogonality. ■

Theorem 3: Let a robust estimator $\mathcal{R}$ be designed by (3.6)–(3.11). When applying it on a system equal to the nominal model, the increased MSE, as compared to the minimum obtainable with a nominal estimator $\mathcal{R}_n$, is

$$\begin{aligned}\mathrm{tr}E(\varepsilon(k)\varepsilon(k)^*) - \mathrm{tr}E(\varepsilon(k)\varepsilon(k)^*)_n &= \\ \|\mathcal{W}(\mathcal{R} - \mathcal{R}_n)z^m D_o^{-1}\boldsymbol{A}_o^{-1}\boldsymbol{N}_o^{-1}\boldsymbol{\beta}_o\|_2^2.\end{aligned} \tag{3.19}$$

Here, $D_o^{-1}\boldsymbol{A}_o^{-1}\boldsymbol{N}_o^{-1}\boldsymbol{\beta}_o$ is the nominal innovations model and $\boldsymbol{\beta}_o$ is obtained from the nominal spectral factorization

$$\boldsymbol{\beta}_o\boldsymbol{\beta}_{o*} = \boldsymbol{N}_o\boldsymbol{B}_o\boldsymbol{C}_o\boldsymbol{C}_{o*}\boldsymbol{B}_{o*}\boldsymbol{N}_{o*} + D_o\boldsymbol{A}_o\boldsymbol{M}_o\boldsymbol{M}_{o*}\boldsymbol{A}_{o*}D_{o*}. \tag{3.20}$$

□

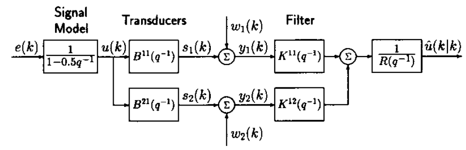


Fig. 2. Model and filter structure in the design example.

Proof: Analogous to that of Theorem 2, by expressing $\mathcal{R}$ as $\mathcal{R}_n + (\mathcal{R} - \mathcal{R}_n)$. ■

Remarks: Expression (3.18) can be used for arbitrary linear estimators $\mathcal{R}_n$, for example, minimax-designs. Expression (3.19) quantifies the price paid in nominal performance for obtaining a robust design. The averaged innovations model in (3.18), and the nominal innovations model in (3.19), together with the filter $\mathcal{W}$, can be seen as weighting functions.

The largest effect of robust filtering is obtained at moderate and high signal-to-noise ratios. If the variance of broad-band measurement noise is increased, the gains of both the nominal and the robust filters decrease. If the noise level is high, performance differences between nominal and robust solutions tend to be small.

IV. A Design Example

Assume that a scalar signal $u(k)$ is to be estimated. It is described by a first order AR-process without uncertainty

$$u(k) = \frac{1}{1 - 0.5q^{-1}}e(k) \quad ; \quad Ee(k)^2 = 1.$$

Thus, $\mathcal{D} = \mathcal{S}/T = 1$, $D_1 = 1$, $D = D_o = 1 - 0.5q^{-1}$, $\hat{\boldsymbol{C}}_o = 1$, and $\hat{C}_1 = 0$. This signal is measured by two transducers ($p = 2$), with nominal models being second order FIR filters. The transducers are modeled by

$$y(k) = (\boldsymbol{B}_o + \boldsymbol{A}_1^{-1}\Delta\boldsymbol{B})u(k) + w(k)$$

with

$$\boldsymbol{B}_o = \begin{pmatrix} B_o^{11} \\ B_o^{21} \end{pmatrix} = \begin{pmatrix} 0.100 + 0.080q^{-2} \\ 1 - 1.4q^{-1} + 0.92q^{-2} \end{pmatrix} \tag{4.1}$$

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 - 0.6q^{-1} \end{pmatrix};$$

$$\Delta\boldsymbol{B} = \begin{pmatrix} \Delta B^{11} \\ \Delta B^{21} \end{pmatrix} = \begin{pmatrix} \Delta b_o^{11} + \Delta b_2^{11}q^{-2} \\ \Delta b_o^{21} + \Delta b_1^{21}q^{-1} + \Delta b_2^{21}q^{-2} \end{pmatrix}.$$

Thus, $\boldsymbol{B}_1 = \boldsymbol{A}_o = \mathbf{I}_2$ and $\boldsymbol{A} = \boldsymbol{A}_1$ are used in (2.5). See Fig. 2.

In the first transducer B^{11}, there is only a single uncertain parameter. It affects the coefficients Δb_o^{11} and Δb_2^{11} with opposite signs, so they have zero mean, variance r_1^2 and cross-covariance $-r_1^2$. In the second transducer, the stochastic coefficients are assumed mutually uncorrelated, with zero means and equal variance r_2^2. Thus, the auto-covariance matrices are

$$\mathbf{P}_{\Delta\mathbf{B}}^{(11,11)} = r_1^2 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}; \quad \mathbf{P}_{\Delta\mathbf{B}}^{(21,21)} = r_2^2 \mathbf{I}_3. \quad (4.2)$$

The scale factors (standard deviations) of the uncertainties are set to

$$r_1 = 0.02; \quad r_2 = 0.10. \quad (4.3)$$

Coefficients of ΔB^{11} and ΔB^{21} are assumed mutually uncorrelated. The complete covariance matrix (2.11) then becomes

$$\mathbf{P}_{\Delta\mathbf{B}} = \begin{pmatrix} \mathbf{P}_{\Delta\mathbf{B}}^{(11,11)} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\Delta\mathbf{B}}^{(21,21)} \end{pmatrix}. \quad (4.4)$$

The measurement noises $w_i(k)$ have variance 0.01. They are white and mutually uncorrelated. Thus, $w(k) = \boldsymbol{M}v(k)$, with

$$\boldsymbol{M} = \boldsymbol{M}_o = 0.1\,\mathbf{I}_2. \quad (4.5)$$

The goal is now to design a filter ($m = 0$), which estimates $u(k)$ based on the two measurements $y_1(k)$ and $y_2(k)$. Frequency weighting is not used used here ($\boldsymbol{W} = 1$), but its influence will be illustrated at the end of the example. In (3.4)–(3.5), we obtain

$$\begin{aligned} \tilde{C}\tilde{C}_* &= \bar{E}(CC_*) = 1 \\ \tilde{\boldsymbol{M}}\tilde{\boldsymbol{M}}_* &= \bar{E}(\boldsymbol{M}\boldsymbol{M}_*) = 0.01\,\mathbf{I}_2 \\ \tilde{\boldsymbol{B}}_C\tilde{\boldsymbol{B}}_{C*} &= \bar{E}(\boldsymbol{B}\tilde{C}\tilde{C}_*\boldsymbol{B}_*) = \hat{\boldsymbol{B}}_o\hat{\boldsymbol{B}}_{o*} + \bar{E}(\Delta\boldsymbol{B}\Delta\boldsymbol{B}_*). \end{aligned} \quad (4.6)$$

Expression (3.7) or (2.10) gives

$$\begin{aligned} \bar{E}(\Delta B^{11}\Delta B_*^{11}) &= \varphi^T \mathbf{P}_{\Delta\mathbf{B}}^{(11,11)} \varphi_*^T \\ &= (1\, q^{-1} q^{-2}) r_1^2 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ q \\ q^2 \end{pmatrix} \\ &= r_1^2(-q^2 + 2 - q^{-2}). \end{aligned} \quad (4.7)$$

Note that $\bar{E}(\Delta B^{11}\Delta B_*^{11})$ has zeros at $z = 1$ and at $z = -1$. Thus, the static gain and the high-frequency gain are assumed to be exactly known. Furthermore

$$\bar{E}(\Delta B^{21}\Delta B_*^{21}) = \varphi^T \mathbf{P}_{\Delta\mathbf{B}}^{(21,21)} \varphi_*^T = 3r_2^2. \quad (4.8)$$

The spectral factorization (3.6) has dimension $p|p = 2|2$. Using (4.6), it reduces to

$$\begin{aligned} \boldsymbol{\beta}\boldsymbol{\beta}_* &= \tilde{\boldsymbol{B}}_C\tilde{\boldsymbol{B}}_{C*} + D\boldsymbol{A}\tilde{\boldsymbol{M}}\tilde{\boldsymbol{M}}_*\boldsymbol{A}_*D_* \\ &= \boldsymbol{A}_1\boldsymbol{B}_o\boldsymbol{B}_{o*}\boldsymbol{A}_{1*} + \bar{E}(\Delta\boldsymbol{B}\Delta\boldsymbol{B}_*) + 0.01D\boldsymbol{A}_1\boldsymbol{A}_{1*}D_* \end{aligned}$$

where

$$\bar{E}(\Delta\boldsymbol{B}\Delta\boldsymbol{B}_*) = \begin{pmatrix} r_1^2(-q^2 + 2 - q^{-2}) & 0 \\ 0 & 3r_2^2 \end{pmatrix}.$$

By using the Newton-based algorithm described in [18], a stable averaged spectral factor, with $\boldsymbol{\beta}(0)$ nonsingular, was found to be (4.9), as shown at the bottom of the page.

In the Diophantine equation (3.11), we use $m = 0$, $\tilde{\boldsymbol{V}} = \boldsymbol{V} = 1$, $\boldsymbol{S} = 1$, $\tilde{C}\tilde{C}_* = 1$, $\boldsymbol{N}_* = \mathbf{I}_2$, $U = 1$, and $T = 1$. Equation (3.11) thus reduces to

$$\boldsymbol{B}_{o*}\boldsymbol{A}_{1*} = \boldsymbol{Q}\boldsymbol{\beta}_* + \boldsymbol{L}_*qD\,\mathbf{I}_2. \quad (4.10)$$

The degrees (3.12) are $nQ = 0, nL_* = 2$. By expressing the polynomial matrices as matrix polynomials, (4.10) becomes

$$\begin{aligned} (\mathbf{B}_0^* + \mathbf{B}_1^*q + \mathbf{B}_2^*q^2)(\mathbf{I}_2 + \mathbf{A}_1^*q) = {} & \mathbf{Q}_0(\boldsymbol{\beta}_0^* + \boldsymbol{\beta}_1^*q + \boldsymbol{\beta}_2^*q^2 + \boldsymbol{\beta}_3^*q^3) \\ & + (\mathbf{L}_0^* + \mathbf{L}_1^*q + \mathbf{L}_2^*q^2) \\ & \times (-0.5 + q)\mathbf{I}_2. \end{aligned}$$

Transpose this equation, and note that since the coefficient matrices are real-valued, $\mathbf{P}_i^{*T} = \mathbf{P}_i$. By equating the two sides for each power of q separately, a linear system of eight equations, in block-Toepliz form, is obtained

$$\begin{pmatrix} \mathbf{B}_0 \\ \mathbf{B}_1 + \mathbf{A}_1\mathbf{B}_0 \\ \mathbf{B}_2 + \mathbf{A}_1\mathbf{B}_1 \\ \mathbf{A}_1\mathbf{B}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta}_0 & -0.5\mathbf{I}_2 & \mathbf{0} & \mathbf{0} \\ \boldsymbol{\beta}_1 & \mathbf{I}_2 & -0.5\mathbf{I}_2 & \mathbf{0} \\ \boldsymbol{\beta}_2 & \mathbf{0} & \mathbf{I}_2 & -0.5\mathbf{I}_2 \\ \boldsymbol{\beta}_3 & \mathbf{0} & \mathbf{0} & \mathbf{I}_2 \end{pmatrix} \begin{pmatrix} \mathbf{Q}_0^T \\ \mathbf{L}_0 \\ \mathbf{L}_1 \\ \mathbf{L}_2 \end{pmatrix}.$$

With numerical values from (4.1) and (4.9), we obtain (x), as found at the bottom of the page. The solution is

$$\begin{aligned} \boldsymbol{Q} &= (0.4005 \quad 0.7746) \\ \boldsymbol{L}_* &= (0.0290 + 0.0200q \quad -0.2053 + 0.3224q - 0.1299q^2). \end{aligned} \quad (4.11)$$

$$\boldsymbol{\beta} = \begin{pmatrix} 0.1339 - 0.01867q^{-1} + 0.01622q^{-2} & 0.07862 - 0.01488q^{-1} + 0.06905q^{-2} \\ -0.1474q^{-1} + 0.2908q^{-2} - 0.1325q^{-3} & 1.1585 - 2.0327q^{-1} + 1.6219q^{-2} - 0.4765q^{-3} \end{pmatrix}. \quad (4.9)$$

$$\begin{pmatrix} .1 \\ 1 \\ 0 \\ -2 \\ .08 \\ 1.76 \\ 0 \\ -.552 \end{pmatrix} = \begin{pmatrix} .1339 & .07862 & -.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.1585 & 0 & -.5 & 0 & 0 & 0 & 0 \\ -.01867 & -.01488 & 1 & 0 & -.5 & 0 & 0 & 0 \\ -.1474 & -2.0327 & 0 & 1 & 0 & -.5 & 0 & 0 \\ .01622 & .06905 & 0 & 0 & 1 & 0 & -.5 & 0 \\ .2908 & 1.6219 & 0 & 0 & 0 & 1 & 0 & -.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -.1325 & -.4765 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} Q_o^{11} \\ Q_o^{12} \\ \ell_o^{11} \\ \ell_o^{12} \\ \ell_1^{11} \\ \ell_1^{12} \\ \ell_2^{11} \\ \ell_2^{12} \end{pmatrix} \quad \text{(x)}$$

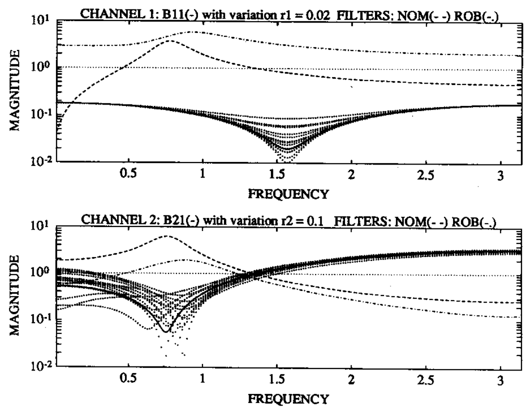


Fig. 3. Bode magnitude plots for the nominal models of the two transducers $B^{11}(q^{-1})$ and $B^{21}(q^{-1})$ (solid). The dotted curves show fifteen realizations of possible true systems. Magnitude plots for the gains from $y_1(k)$ (upper) and $y_2(k)$ (lower), are shown for the robust estimator (dash-dotted) and for the nominal Wiener filter design (dashed).

Finally, the robust estimator (3.10) becomes

$$\mathcal{R} = Q\beta^{-1}A_1 = \frac{1}{R_r}(K_r^{11}\ K_r^{12}) \tag{4.12}$$

where the monic denominator is $R_r(q^{-1}) = \det \beta\, (q^{-1})/\det \beta_0$. We obtain

$$K_r^{11} = 2.9922 - 4.5138q^{-1} + 2.7365q^{-2} - 0.5687q^{-3}$$
$$K_r^{12} = 0.4655 - 0.3341q^{-1} - 0.06445q^{-2} + 0.05841q^{-3}$$
$$R_r = 1 - 1.8193q^{-1} + 1.6043q^{-2} - 0.6584q^{-3} + 0.08479q^{-4} + 0.009182q^{-5}.$$

A corresponding nominal estimator (with no uncertainty assumed) is given by $(K_n^{11}\ K_n^{12})/R_n$, with

$$K_n^{11} = 0.7419 - 1.0943q^{-1} + 0.3617q^{-2}$$
$$K_n^{12} = 0.8792 - 0.3767q^{-1} - 0.03145q^{-2}$$
$$R_n = 1 - 1.7786q^{-1} + 1.4269q^{-2} - 0.3938q^{-3}.$$

Fig. 3 shows the Bode magnitude plots for the robust and nominal estimators. Also shown is the nominal transducer model and 15 randomly chosen systems. These were generated by using $B = B_o + A_1^{-1}\Delta B$, with covariance matrix (4.4) and Gaussian distributions. The channel B^{11} has its uncertainty concentrated around the notch, while B^{21} is uncertain mainly at low frequencies.

The gains of the nominal estimator (dashed curves) are determined exclusively by the nominal signal to noise ratios. The gains of the robust estimator (dash-dotted) are determined by the balance between noise levels and model uncertainties in the two channels. For example, the robust filter "knows" that channel 1 is well known, as compared to channel 2. Consequently, a higher gain is used from $y_1(k)$ as compared to the nominal case, and a lower gain from $y_2(k)$. The difference, as compared to nominal design, is largest at low frequencies. There, the dynamics of channel 1 is almost perfectly known, while channel 2 is very uncertain. The nominal filter gain in channel 2 is an approximate inverse of the nominal transducer. In contrast to the nominal filter, the robust filter has hardly any peak at the (uncertain) notch around $\omega = 0.7$. It utilizes channel 1 more at this frequency.

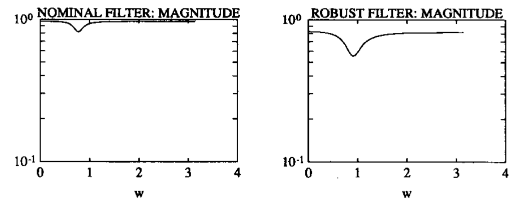


Fig. 4. Bode magnitude plots for the transfer function from $u(k)$ to $\hat{u}(k|k)$for the nominal system, with robust and nominal estimators.

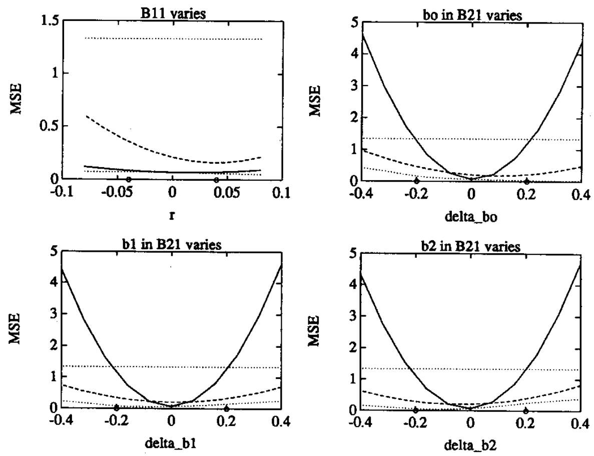


Fig. 5. MSE for robust (dashed) and nominal filter (solid). One of the four uncertain parameters is varied, while the others are held at nominal values. Also shown is the variance of $u(k)$ (upper dotted). It corresponds to the error caused by the "trivial" estimate $\hat{u}(k) = 0$. The lower dotted curve is the lower bound, achievable with knowledge of the true parameter values. Rings (o) indicate the two standard deviation limits of each parameter.

Fig. 4 shows Bode magnitude plots of transfer functions from $u(k)$ to $\hat{u}(k|k)$. Since the noise levels are rather low, the nominal estimator performs an almost complete inversion of the nominal transducers. The robust estimator is somewhat more cautious, but it also accomplishes a rather good inversion. It utilizes the two measurement signals differently than the nominal estimator.

Fig. 5 shows the mean square estimation error, when one of the uncertain parameters

$$\Delta b_o^{11} = -\Delta b_2^{11} \triangleq r; \qquad \Delta b_o^{21}; \qquad \Delta b_1^{21}; \qquad \Delta b_2^{21}$$

is varied, while the others are zero. The four parameters above span the set of assumed true systems, the extended design model. On average, over the four uncorrelated stochastic coefficients, the MSE is 0.32 for the robust filter and 0.90 for the nominal design. Note, however, that when r is varied, the robust design (dashed) is actually slightly more sensitive than the nominal design. This is a price paid for reducing the sensitivity in the other dimensions.[6]

[6] It is natural that the robust filter has a somewhat increased sensitivity to model errors in channel 1; it has higher gain in that channel. This result is due to the much larger uncertainty in channel 2 at most frequencies. A designer worried about this effect could simply increase the value of the standard deviation r_1 used in the design.

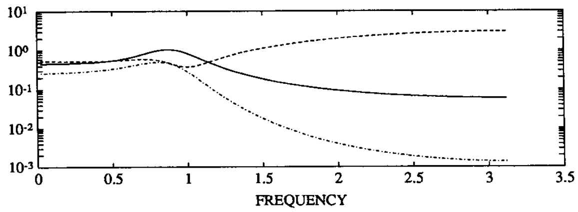


Fig. 6. The average spectral density of the estimation error $u(k) - \hat{u}(k|k)$. Shown in the plot are the lower bound, according to (3.15), (dash-dotted) and the average spectral densities obtained with (dashed) and without (solid) frequency weighting.

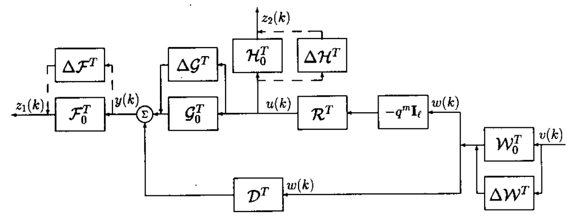


Fig. 7. A block diagram which is dual to the one in Fig. 1. Arrows are reversed, summation points and node points are interchanged, and transfer functions are transposed. The corresponding robust control problem is to design $\mathcal{R}^T$ to minimize the average, over the class of models, of the $\mathcal{H}_2$-norm of the transfer function from v to $z = (z_1^T \; z_2^T)^T$.

The robust estimator, of course, does not perform as well as the nominal one in the nominal case. This is mainly due to its somewhat lower gain from $u(k)$ to $\hat{u}(k|k)$; see Fig. 4. It is evident from Fig. 5 that this performance loss is very small, as compared to the improvement in nonideal situations.[7]

As an alternative, we tried to investigate minimax designs, i.e., worst case $\mathcal{H}_2$-designs, assuming rectangular parameter distributions. This turned out to be prohibitively difficult, since no point where $\min_{\mathcal{R}} \max_{\Delta B} \varepsilon(k)^2 = \max_{\Delta B} \min_{\mathcal{R}} \varepsilon(k)^2$ could be found. These difficulties were in marked contrast to the ease of designing a cautious Wiener filter based on the averaged $\mathcal{H}_2$-criterion.

Let us finally illustrate the effect of using a frequency-dependent weighting function in the criterion (2.3). Assume that the performance at frequencies close to $\omega = 0.9$ is of particular importance. The choice

$$\mathcal{W} = \frac{V}{U} = \frac{1}{1 - 1.2184q^{-1} + 0.9604q^{-2}} \tag{4.13}$$

with a high resonance peak ($|z| = 0.98$) at $\omega = 0.9$ should, according to Corollary 1, result in a performance, at that frequency, close to the lower bound. Fig. 6 confirms this. Performance is substantially degraded at higher frequencies, however, where estimation accuracy is not emphasized.

V. Robust Feedforward Control

A class of feedforward control problems turns out to be dual to the filtering problems discussed in Section II. We include a brief separate discussion of them, since it offers several engineering insights. Feedforward compensation does not affect the classical sensitivity function. The effect of an $x\%$ model deviation at a particular frequency, however, on e.g., the step response, will very much depend on the (nominal) magnitude of the transfer function at that frequency. As the gain at a particular frequency is increased by a feedforward link, model errors at that frequency become more and more noticeable. Therefore, it is of value to take model uncertainty into account explicitly in the feedforward design.

[7]It can also be noted that it is of advantage to use both channels. The minimal MSE, for channels equal to the nominal models, is 0.07 if both channels are used. It is 0.59 if only channel 1 is used and 0.11 if only channel 2 is used. The average MSE of the robust filter (0.32) is in fact lower than the nominal MSE for an estimator which uses only channel 1 (0.59).

To stress the duality to filtering problems, the output of an uncertain but stable model will be described by

$$\begin{aligned} y(k) &= (\mathcal{G}_o^T + \Delta\mathcal{G}^T)u(k) + \mathcal{D}^T w(k) \\ &= (\boldsymbol{B}_o^T \boldsymbol{A}_o^{-T} + \Delta\boldsymbol{B}^T \boldsymbol{B}_1^T \boldsymbol{A}_1^{-T})u(k) + \frac{1}{T}\boldsymbol{S}^T w(k) \end{aligned} \tag{5.1}$$

where $\boldsymbol{A}^{-T}$ denotes inverse and transpose. The rational and polynomial matrices above have properties as outlined in Section II. The matrices $\mathcal{G}^T$ and $\mathcal{D}^T$ may contain delays. Based on possibly delayed or advanced measurements of

$$\begin{aligned} w(k) &= (\mathcal{W}_o^T + \Delta\mathcal{W}^T)v(k) \\ &= \left(\frac{1}{U_o}\boldsymbol{V}_o^T + \frac{1}{U_1}\Delta\boldsymbol{V}^T\boldsymbol{V}_1^T\right)v(k); \\ E(v(k)^T v(k)) &= \mathbf{I}_\ell \end{aligned} \tag{5.2}$$

a stable controller

$$u(k) = -\mathcal{R}^T w(k+m) \tag{5.3}$$

is to be designed to minimize the averaged $\mathcal{H}_2$-norm of the transfer function from v to $(z_1^T \; z_2^T)^T \triangleq ((\mathcal{F}^T y)^T \; (\mathcal{H}^T u)^T)^T$

$$J' = \bar{E}\left\| \begin{matrix} \mathcal{F}^T(\mathcal{D}^T - \mathcal{G}^T\mathcal{R}^T z^m)\mathcal{W}^T \\ -\mathcal{H}^T\mathcal{R}^T z^m \mathcal{W}^T \end{matrix} \right\|_2^2. \tag{5.4}$$

The control weighting $\mathcal{H}^T = \boldsymbol{M}^T\boldsymbol{N}^{-T}$ and the output weighting $\mathcal{F}^T = \boldsymbol{C}^T/D$ are normally specified by the designer and exactly known, i.e., $\Delta\mathcal{H}^T = 0$, $\Delta\mathcal{F}^T = 0$.

Problem formulation (5.1)–(5.4) may represent a disturbance measurement feedforward design. When $m > 0$, the disturbance $w(k)$ can be measured before it affects the system via $\mathcal{D}^T$. (Such a situation could, equivalently, be described by a delay q^{-m} in $\mathcal{D}^T$.) The formulation also covers reference feedforward problems, feedforward decoupling, and model matching. Then, $w(k)$ is a command signal and $\mathcal{W}^T v(k)$ is a (possibly uncertain) stochastic model, describing its second order properties. A servo filter $\mathcal{R}^T$ is to be designed, so that the output $-\mathcal{G}^T u(k)$ optimally follows the response model $\mathcal{D}^T w(k)$.[8] In decoupling problems, $\mathcal{D}^T$ is diagonal.

The duality of feedforward control to the previously discussed filtering problem has been described in [5] for the

[8]The corresponding nominal result (without uncertainty) was discussed in [35] and in [5]. See also [27], where common denominator forms were used. The robust controller for SISO systems was presented in [36]. In [6], the applicability of the averaged $\mathcal{H}_2$-criterion is demonstrated for a somewhat more general class of open-loop type problems.

nominal case. The corresponding result in the presence of model errors is given below.

Theorem 4: A feedforward filter, which solves the robust control problem (5.1)–(5.4) under assumptions A1–A4 is given by transposing $\mathcal{R}$ from (3.10), or

$$u(k) = -A^T N^T \beta^{-T} Q^T \tilde{V}^{-T} \frac{1}{T} w(k+m) \quad (5.5)$$

where β, Q, and $\tilde{V}$ are given by (3.6), (3.11), and (3.9), respectively. □

Proof: The averaged $\mathcal{H}_2$-norm is invariant under transposition. Thus, it fulfills the basic requirement of [5]. By extending the discussion in [5] to uncertain models the result is obtained. ■

The uncertainty $\Delta\mathcal{W}^T$ of the disturbance or reference model (5.2) enters via the spectral factorization (3.9). Transposition of (3.9) gives

$$\tilde{V}^T \tilde{V}_*^T = U_1 V_o^T V_{o*}^T U_{1*} + U_o V_1^T \bar{E}(\Delta V^T \Delta V_*^T) V_{1*}^T U_{o*}. \quad (5.6)$$

This is the kind of left spectral factorization encountered when two noise sources are described by one innovations model. In fact, the uncertainty $\Delta\mathcal{W}^T$ has exactly the same effect on the controller design as would a measurement noise on $w(k)$, with spectral density $V_1^T \bar{E}(\Delta V^T \Delta V_*^T) V_{1*}^T / U_1 U_{1*}$. We do not need to solve a right spectral factorization (3.9) in this problem. The left spectral factorization (5.6) can be solved instead.

As in the dual filtering case, uncertainty in the direct feedthrough, or response model $\mathcal{D}^T = \mathcal{S}^T/T$ does not affect the optimal solution, if it is independent of the uncertainties in $\mathcal{G}^T$.

VI. Conclusions

A method for designing robust filters and feedforward controllers, based on imperfectly known linear models, has been presented. Modeling errors were described by sets of models, parameterized by random variables with known covariances. A robust design was obtained by minimizing the $\mathcal{H}_2$-norm, averaged with respect to the assumed model errors. The estimator minimizes this criterion by balancing model uncertainties against noise properties, at different frequencies and in different measurement channels. When using robust filtering, the greatest sensitivity reduction is obtained at moderate and high signal to noise ratios. Dually, the largest impact of robust control is obtained for designs with low input penalties.

One variant of the discussed filtering problems is to explicitly define a part of the measurement vector as being a noise-free signal. This signal could e.g., represent known inputs to the system. Such a formulation is also of use in the optimization of decision feedback equalizers for digital communications [33], [34].

There exist efficient numerical algorithms based on a polynomial equations approach. (We have implemented them as MATLAB .m-files, and the code is available upon request.) For multivariable problems of high order and high signal vector dimension p, Riccati-based algorithms do, however, perform better numerically. For dimensions of the measurement vector up to, say, four, algorithms based directly on polynomial manipulation can in general be used safely. For higher dimensions, we recommend analytical solutions to be obtained by an input–output approach, to gain engineering insight, but algorithms e.g., for spectral factorization to be based on state space formulations, cf. [22]

Appendix

A. Proofs of Lemmas

Proof of Lemma 1: Let $G = [G_1^T \cdots G_n^T]^T$, where $G_1 \cdots G_n$ represent the n polynomial row vectors of the $n|m$ matrix. Then, with H of dimension $m|m$, the ijth element of $\bar{E}(GHG_*)$ can be expressed as

$$\begin{aligned}[\bar{E}(GHG_*)]_{ij} &= \bar{E}(G_i H G_{j*}) \\ &= \bar{E}(\text{tr} G_{j*} G_i H) = \text{tr}\bar{E}(G_{j*}G_i)\bar{E}(H).\end{aligned}$$

In the last equality, we used the fact that all elements of $G_{j*}G_i$ are independent of all those of H. We also have that

$$\text{tr}\bar{E}(G_{j*}G_i)\bar{E}(H) = \bar{E}(G_i \bar{E}(H) G_{j*}) = [\bar{E}(G\bar{E}(H)G_*)]_{ij}$$

which proves (3.3), since $\bar{E}(\cdot)$ operates on all elements of GHG_*. ■

Proof of Lemma 2: If the coefficients of the elements of a polynomial matrix $\Delta P(q^{-1})$ are stochastic variables, then so are the coefficients of the elements in $\Delta P \Delta P_*$. The coefficients of the polynomial elements in ΔC and ΔB are independent and so are the coefficients of the elements in $\Delta C \Delta C_*$ and $\Delta B \Delta B_*$. By defining CC_* as H and ΔB as G in Lemma 1 and using (3.4), the right-hand side of (3.1) becomes

$$\begin{aligned}&N\bar{E}[BCC_*B_*]N_* + DA\bar{E}(MM_*)A_*D_* = \\ &= N\bar{E}[B\bar{E}(CC_*)B_*]N_* + DA\bar{E}(MM_*)A_*D_* \\ &= N\bar{E}[B\tilde{C}\tilde{C}_*B_*]N_* + DA\tilde{M}\tilde{M}_*A_*D_*.\end{aligned}$$

By once more utilizing (3.4), we obtain (3.6).

B. Calculation of Averaged Polynomial Matrices

Consider the matrices discussed in Lemma 1. Let the polynomial matrices $G = \Delta P$ and H be of dimensions $n|m$ and $m|m$, respectively. Denote the i-th row of ΔP by

$$\Delta P_i = [\Delta P^{i1} \cdots \Delta P^{im}]$$

where ΔP^{ij} are polynomials with stochastic coefficients. If H is assumed deterministic, the ijth element of $\bar{E}(\Delta P H \Delta P_*)$ can be written as

$$\begin{aligned}\bar{E}[\Delta P H \Delta P_*]_{ij} &= \text{tr}\bar{E}(\Delta P_{j*}\Delta P_i)H = \\ &= \text{tr} H \bar{E}[\Delta P_*^{j1} \cdots \Delta P_*^{jm}]^T [\Delta P^{i1} \cdots \Delta P^{im}] \\ &= \text{tr} H \bar{E} \begin{bmatrix} \Delta P^{i1}\Delta P_*^{j1} & \cdots & \Delta P^{im}\Delta P_*^{j1} \\ \vdots & \ddots & \vdots \\ \Delta P^{i1}\Delta P_*^{jm} & \cdots & \Delta P^{im}\Delta P_*^{jm} \end{bmatrix}.\end{aligned}$$

By using (2.10), we readily obtain (3.7) . ■

C. Proof of Theorem 1

A technique for constructive derivation of polynomial design equations for Wiener filters was presented in [1]. This method is utilized here to minimize (2.3). The estimation error is given by

$$\varepsilon(k) = \mathcal{W}(f(k) - \hat{f}(k|k+m)). \tag{C.1}$$

All admissible alternatives to a proposed estimate $\hat{f}(k|k+m)$, can be described by

$$\hat{d}(k) = \mathcal{R}y(k+m) + \tilde{\nu}(k); \qquad \tilde{\nu}(k) = \mathcal{M}y(k+m). \tag{C.2}$$

Here, $\mathcal{M}$ is a rational, stable and causal, but otherwise arbitrary transfer function. Define the weighted variation

$$\begin{aligned}\nu(k) &\triangleq \mathcal{W}\tilde{\nu}(k)\\ &= \frac{1}{U}\boldsymbol{V}\mathcal{M}q^m\left(\boldsymbol{A}^{-1}\boldsymbol{B}\frac{1}{D}\boldsymbol{C}e(k) + \boldsymbol{N}^{-1}\boldsymbol{M}v(k)\right).\end{aligned}$$

Optimality of (2.2) is obtained if no perturbation $\nu(k)$ will improve the average estimator performance. This occurs if and only if the error $\varepsilon(k)$ is orthogonal to any admissible weighted estimator variation $\nu(k)$. In other words

$$\text{tr}\bar{E}E(\varepsilon(k)\nu(k)^*) = \text{tr}\bar{E}E(\varepsilon(k)^*\nu(k)) = 0.$$

Then, the perturbed criterion value becomes

$$\begin{aligned}\bar{J} &= \text{tr}\bar{E}E[\mathcal{W}(f(k) - \hat{d}(k))][\mathcal{W}(f(k) - \hat{d}(k))]^*\\ &= \text{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^* - \varepsilon(k)\nu(k)^*\\ &\quad - \varepsilon(k)^*\nu(k) + \nu(k)\nu(k)^*)\\ &= \text{tr}\bar{E}E(\varepsilon(k)\varepsilon(k)^* + \nu(k)\nu(k)^*).\end{aligned} \tag{C.3}$$

This expression is evidently minimized by $\nu(k) = 0$.

Since all transfer functions in (2.4) are assumed stable, both $\varepsilon(k)$ and $\nu(k)$ are stationary. Parseval's formula may then be used to express $\text{tr}\bar{E}E\varepsilon(k)\nu(k)^*$ as

$$\begin{aligned}&\text{tr}\bar{E}E\frac{1}{U}\boldsymbol{V}\left\{\left(\frac{1}{T}\boldsymbol{S} - q^m\mathcal{R}\boldsymbol{A}^{-1}\boldsymbol{B}\right)\frac{1}{D}\boldsymbol{C}e(k) - q^m\mathcal{R}\boldsymbol{N}^{-1}\boldsymbol{M}v(k)\right\}\\ &\times\left\{\frac{1}{U}\boldsymbol{V}\mathcal{M}q^m\left(\boldsymbol{A}^{-1}\boldsymbol{B}\frac{1}{D}\boldsymbol{C}e(k) + \boldsymbol{N}^{-1}\boldsymbol{M}v(k)\right)\right\}^*\\ &= \text{tr}\bar{E}\frac{1}{2\pi j}\oint_{|z|=1}\frac{1}{UU_*DD_*}\\ &\times\boldsymbol{V}\left\{z^{-m}\frac{1}{T}\boldsymbol{S}\boldsymbol{C}\boldsymbol{C}_*\boldsymbol{B}_*\boldsymbol{A}_*^{-1} - \mathcal{R}\boldsymbol{A}^{-1}\boldsymbol{N}^{-1}\right.\\ &\times\left.(\boldsymbol{N}\boldsymbol{B}\boldsymbol{C}\boldsymbol{C}_*\boldsymbol{B}_*\boldsymbol{N}_* + D\boldsymbol{A}\boldsymbol{M}\boldsymbol{M}_*\boldsymbol{A}_*D_*)\boldsymbol{N}_*^{-1}\boldsymbol{A}_*^{-1}\right\}\\ &\times\mathcal{M}_*\boldsymbol{V}_*\frac{dz}{z}.\end{aligned} \tag{C.4}$$

Note that $\boldsymbol{A}$ and $\boldsymbol{N}$ commute, since they are diagonal. We are allowed to move the expectation $\bar{E}$ inside the integration, since, for any particular realization of the elements of $\Delta\boldsymbol{C}$, $\Delta\boldsymbol{B}$, $\Delta\boldsymbol{M}$, and $\Delta\boldsymbol{V}$, all elements of the integrand are Riemann integrable on the unit circle; see e.g., [17, Theorem 3.8]. The use of the trace rotation, $\text{tr}\boldsymbol{V}\{\cdots\}\mathcal{M}_*\boldsymbol{V}_* = \text{tr}\boldsymbol{V}_*\boldsymbol{V}\{\cdots\}\mathcal{M}_*$, the spectral factorizations (3.1) and (3.9) and the Assumption A3, leads to

$$\begin{aligned}\text{tr}\bar{E}E\varepsilon(k)\nu(k)^* &= \text{tr}\frac{1}{2\pi j}\oint\frac{1}{UU_*DD_*}\tilde{\boldsymbol{V}}_*\tilde{\boldsymbol{V}}\\ &\times\left\{z^{-m}\bar{E}(\frac{1}{T}\boldsymbol{S}\boldsymbol{C}\boldsymbol{C}_*\boldsymbol{B}_*)\boldsymbol{A}_*^{-1}\right.\\ &\left.-\mathcal{R}\boldsymbol{A}^{-1}\boldsymbol{N}^{-1}\boldsymbol{\beta}\boldsymbol{\beta}_*\boldsymbol{N}_*^{-1}\boldsymbol{A}_*^{-1}\right\}\mathcal{M}_*\frac{dz}{z}.\end{aligned} \tag{C.5}$$

From (C.5) it is now easy to see that uncertainties in

$$\mathcal{D} = \frac{1}{T_o}\boldsymbol{S}_o + \frac{1}{T_1}\boldsymbol{S}_1\Delta\boldsymbol{S} = \frac{1}{T_oT_1}\left(\hat{\boldsymbol{S}}_o + \hat{\boldsymbol{S}}_1\Delta\boldsymbol{S}\right) \triangleq \frac{1}{T}\boldsymbol{S} \tag{C.6}$$

independent of $\Delta\boldsymbol{C}$, $\Delta\boldsymbol{B}$ and $\Delta\boldsymbol{V}$, will not affect the filter design since, using Assumption A1

$$\bar{E}(\frac{1}{T}\boldsymbol{S}\boldsymbol{C}\boldsymbol{C}_*\boldsymbol{B}_*) = \bar{E}(\frac{1}{T}\boldsymbol{S})\bar{E}(\boldsymbol{C}\boldsymbol{C}_*)\bar{E}(\boldsymbol{B}_*) = \frac{1}{T_o}\boldsymbol{S}_o\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}. \tag{C.7}$$

(The minimal criterion value will be affected, however, which is evident form (C.12), below.) In the sequel we thus use $T_o = T$ and $\boldsymbol{S}_o = \boldsymbol{S}$.

Using (C.7) in (C.5) and extracting T to the left and $\boldsymbol{A}_*^{-1}$ to the right now gives

$$\begin{aligned}\text{tr}\bar{E}E\varepsilon(k)\nu(k)^* &= \text{tr}\frac{1}{2\pi j}\oint\frac{1}{UU_*TDD_*}\tilde{\boldsymbol{V}}_*\tilde{\boldsymbol{V}}\\ &\left\{z^{-m}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*} - T\mathcal{R}\boldsymbol{A}^{-1}\boldsymbol{N}^{-1}\boldsymbol{\beta}\boldsymbol{\beta}_*\boldsymbol{N}_*^{-1}\right\}\\ &\times\boldsymbol{A}_*^{-1}\mathcal{M}_*\frac{dz}{z}.\end{aligned} \tag{C.8}$$

To make (C.8) zero, all poles inside $|z| = 1$ are eliminated. This is achieved if, in every element of the integrand, all such poles are cancelled by zeros. We first cancel what can be cancelled by means of $\mathcal{R}$ directly. Thus

$$\mathcal{R} = \frac{1}{T}\tilde{\boldsymbol{V}}^{-1}\boldsymbol{Q}\boldsymbol{\beta}^{-1}\boldsymbol{N}\boldsymbol{A} \tag{C.9}$$

where $\boldsymbol{Q}(z^{-1})$ is undetermined. Inserting (C.9) into (C.8) gives $\text{tr}\bar{E}E\varepsilon(k)\nu(k)^*$ as

$$\begin{aligned}&\text{tr}\frac{1}{2\pi j}\oint\frac{1}{UU_*TDD_*}\tilde{\boldsymbol{V}}_*\\ &\left\{z^{-m}\tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_* - \boldsymbol{Q}\boldsymbol{\beta}_*\right\}\boldsymbol{N}_*^{-1}\boldsymbol{A}_*^{-1}\mathcal{M}_*\frac{dz}{z}.\end{aligned} \tag{C.10}$$

Now, U, T, D, $\boldsymbol{A}$, $\boldsymbol{N}$ are all stable, so they have zeros only inside $|z| = 1$. The poles of $\mathcal{M}$ are inside $|z| = 1$. This means that U_*^{-1}, D_*^{-1}, $\boldsymbol{A}_*^{-1}$, $\boldsymbol{N}_*^{-1}$, and $\mathcal{M}_*$ have all their poles outside $|z| = 1$. No poles will thus exist inside $|z| = 1$ in (C.10), if and only if

$$z^{-m}\tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_* - \boldsymbol{Q}\boldsymbol{\beta}_* = z\boldsymbol{L}_*UTD\,\mathbf{I}_p \tag{C.11}$$

for some polynomial matrix $\boldsymbol{L}_*(z)$. This is (3.11), if q is substituted for z. The filter (C.9) coincides with (3.10). Necessity follows because choices of $\mathcal{R}$ other that (C.9) correspond to $\nu(k) \neq 0$ in (C.3).

Unique solvability of (3.11) is demonstrated as follows. The Diophantine equation will always have one or several solutions, since the invariant polynomials of $UTD\,\mathbf{I}_p$ are all stable, while those of $\boldsymbol{\beta}_*$ are all unstable. Thus, there exist no common invariant factors. Let $(\boldsymbol{Q}_0, \boldsymbol{L}_{0*})$ be one solution pair. Every solution to (3.11) can then be expressed as

$$(\boldsymbol{Q}, \boldsymbol{L}_*) = (\boldsymbol{Q}_0 - \boldsymbol{X} qUTD\mathbf{I}_p \ , \ \boldsymbol{L}_{0*} + \boldsymbol{X}\boldsymbol{\beta}_*)$$

where the polynomial matrix $\boldsymbol{X}(q, q^{-1})$ is undetermined. Now, $\boldsymbol{Q}$ is required to be causal, so it can not have any positive powers of q as arguments, while $\boldsymbol{L}_*$ must contain no negative powers of q, to assure optimality. Thus, $\boldsymbol{X}(q, q^{-1}) = 0$ is the only choice. We conclude that the solution to (3.11) is unique.

The degrees (3.12) are determined by the requirement that the maximum powers of q^{-1} and q are covered on both sides of (3.11). They assure that the number of unknowns equal the number of equations in the corresponding linear system of equations. For details, see [1] or [3].

The minimal average estimation error, $J_{min} = \mathrm{tr}\bar{E}E\ (\varepsilon(k)\ \varepsilon(k)^*)_{min}$, is obtained as follows. First insert (3.10) into the criterion (2.3), use Parseval's formula, take expectation and use (3.1), (3.9) and (3.5) in this order. Then we obtain

$$\begin{aligned} J_{min} = \mathrm{tr}\frac{1}{2\pi j} \oint_{|z|=1} \frac{1}{UU_*TT_*DD_*} \tilde{\boldsymbol{V}}_* \tilde{\boldsymbol{V}} \Big\{ & \boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\boldsymbol{S}_* \\ & - \boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{A}_*^{-1}\boldsymbol{\mathcal{R}}_*T_*z^{-m} - z^m T\boldsymbol{\mathcal{R}}\boldsymbol{A}^{-1}\hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\boldsymbol{S}_* \\ & + T\boldsymbol{\mathcal{R}}\boldsymbol{A}^{-1}\boldsymbol{N}^{-1}\boldsymbol{\beta}\boldsymbol{\beta}_*\boldsymbol{N}_*^{-1}\boldsymbol{A}_*^{-1}\boldsymbol{\mathcal{R}}_*T_* \Big\} \frac{dz}{z}. \end{aligned} \quad \text{(C.12)}$$

Now, the use of (C.9), $\mathrm{tr}\tilde{\boldsymbol{V}}_*\tilde{\boldsymbol{V}}\{\cdots\} = \mathrm{tr}\tilde{\boldsymbol{V}}\{\cdots\}\tilde{\boldsymbol{V}}_*$ and completing the square gives

$$\begin{aligned} J_{min} = \mathrm{tr}\frac{1}{2\pi j} \oint_{|z|=1} \frac{1}{UU_*TT_*DD_*} \Big\{ & \tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\boldsymbol{S}_*\tilde{\boldsymbol{V}}_* \\ & + \left(z^{-m}\tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1} - \boldsymbol{Q}\right) \\ & \times \left(\boldsymbol{\beta}^{-1}\boldsymbol{N}\hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\boldsymbol{S}_*\tilde{\boldsymbol{V}}_*z^m - \boldsymbol{Q}_*\right) \\ & - \tilde{\boldsymbol{V}}\boldsymbol{S}\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\hat{\boldsymbol{B}}_{o*}\boldsymbol{N}_*\boldsymbol{\beta}_*^{-1}\boldsymbol{\beta}^{-1}\boldsymbol{N}\hat{\boldsymbol{B}}_o\tilde{\boldsymbol{C}}\tilde{\boldsymbol{C}}_*\boldsymbol{S}_*\tilde{\boldsymbol{V}}_* \Big\} \frac{dz}{z}. \end{aligned}$$

Finally, use Diophantine equation (C.11) in the middle term and rearrange the terms to obtain expression (3.13). ∎

REFERENCES

[1] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations," *IEEE Trans. Signal Processing*, vol. 39, pp. 2387–2399, 1991.

[2] ——, "Optimal filtering problems," in *Polynomial Methods in Optimal Control and Filtering*, K. Hunt, Ed. London: Peter Peregrinus, 1993.

[3] ——, "Derivation and design of Wiener filters using polynomial equations," *Control and Dynamic Systems, vol 64: Stochastic Techniques in Digital Signal Processing Systems*, C. T. Leondes, Ed. New York: Academic, 1994, pp. 353–418.

[4] P. Balaban and J. Salz, "Optimum diversity combining and equalization in digital data transmission with applications to cellular mobile radio—Part I: Theoretical considerations," *IEEE Trans. Communications*, vol. 40, pp. 885–894, 1992.

[5] B. Bernhardsson and M. Sternad, "Feedforward control is dual to deconvolution," *Int. J. Contr.*, vol. 57, pp. 393–405, 1993.

[6] B. Bernhardsson, "Robust stochastic performance optimization," IFAC World Congress, Sydney, Australia, July 1993, vol. 8, pp. 67–70. Also in *Topics in Digital and Robust Control of Linear Systems*. Ph.D. dissertation, TFRT-1039-SE, Dept. of Automatic Control, Lund Institute of Technology, Lund, Sweden, 1992.

[7] P. Bolzern, P. Colaneri, and G De Nicolao, "Optimal robust filtering for linear systems subject to time-varying parameter perturbations," in *Proc. 32nd Conf. Decis. Contr.*, San Antonio, TX, Dec. 1993, pp. 1018–1023.

[8] B. Carlsson, A. Ahlén, and M. Sternad, "Optimal differentiation based on stochastic signal models," *IEEE Trans. Signal Processing*, vol 39, pp. 341–353, 1991.

[9] R. C. Chung and P. R. Bélanger, "Minimum-sensitivity filter for linear time-invariant stochastic systems with uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. 21, pp. 98–100, 1976.

[10] J. A. D'Appolito and C. E. Hutchinson, "A minimax approach to the design of low sensitivity state estimators," *Automatica*, vol. 8, pp. 599–608, 1972.

[11] G. C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with application to model order selection," *IEEE Trans. Automat. Contr.*, vol. 27, pp. 913–928, 1992.

[12] G. C. Goodwin and M. E. Salgado, "A stochastic embedding approach for quantifying uncertainty in the estimation of restricted complexity models," *Int. J. Adaptive Contr. Signal Processing*, vol. 3, pp. 333–356, 1989.

[13] M. J. Grimble, "Wiener and Kalman filters for systems with random parameters," *IEEE Trans. Automat. Contr.*, vol. 29, pp. 552–554, 1984.

[14] M. J. Grimble and A. ElSayed, "Solution to the $\mathcal{H}_\infty$ optimal linear filtering problem for discrete-time systems," *IEEE Trans. Signal Processing*, vol. 38, pp. 1092–1104, 1990.

[15] W. F. Haddad and D. S. Bernstein, "The optimal projection equations for reduced-order discrete-time state estimation for linear systems with multiplicative noise," *Syst. Contr. Lett.*, vol 8, pp. 381-388, 1987.

[16] ——, "Robust, reduced-order, nonstrictly proper state estimation via the optimal projection equations with guaranteed cost bounds," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 591–595, 1988.

[17] A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* New York: Academic, 1970.

[18] J. Ježek and V. Kučera, "Efficient algorithm for matrix spectral factorization," *Automatica*, vol. 21, pp. 663–669 1985.

[19] S. A. Kassam and T. L. Lim, "Robust Wiener filters," *J. Franklin Inst.*, pp. 171–185, 1977.

[20] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: a survey," *Proc. IEEE*, vol. 73, pp. 433–481, 1985.

[21] V. Kučera, *Discrete Linear Control. The Polynomial Equations Approach.* New York: Wiley, 1979.

[22] ——, "Factorization of rational spectral matrices: a survey," in *Proc. IEE Control'91*, Edinburgh, Mar. 1991, pp. 1074–1078.

[23] W. Lee, *Mobile Cellular Telecommunications Systems.* New York: McGraw-Hill, 1989.

[24] C. T. Leondes and J. O. Pearson, "A minimax filter for systems with large plant uncertainties," *IEEE Trans. Automat. Contr.*, vol. 17, pp. 266–268, 1972.

[25] E. Lindskog, M. Sternad, and A. Ahlén, "Designing decision feedback equalizers to be robust with respect to channel time variations," in *Proc. Nordic Radio Society Symp. Disturbance Resistant Radio and Radar*, Uppsala, Sweden, Nov 10–11 1993.

[26] C. J. Martin and M. Mintz, "Robust filtering and prediction for linear systems with uncertain dynamics: a game-theoretic approach," *IEEE Trans. Automat. Contr.*, vol. 28, pp. 888–896, 1983.

[27] R. H. Milocco, "Optimal design of feedforward regulators for multivariable systems," in *Proc. Automat. Contr. Conf.*, Chicago, IL, June 24–26, 1992, pp. 780–784.

[28] G. Moustakides and S. A. Kassam, "Robust Wiener filters for random signals in correlated noise," *IEEE Trans. Inform. Theory*, vol. 29, pp. 614–619, 1983.

[29] K. Öhrn, "Design of multivariable cautious Wiener filters: A probabilistic approach," Licentiate thesis, Dept. Technology, Uppsala University, Uppsala, Sweden, in preparation.

[30] I. R. Petersen and D. C. McFarlane, "Robust state estimation for uncertain systems," in *Proc. 30th Conf. Decis. Contr.*, Brighton, England, Dec. 1991, pp. 2630–2631.

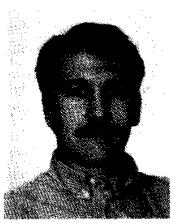
[31] H. V. Poor, "On robust Wiener filtering," *IEEE Trans. Automat. Contr.*, vol. 25, pp. 531–536, 1980.

[32] J. L. Speyer and D. E. Gustafson, "An approximation method for estimation in linear systems with parameter uncertainty," *IEEE Trans. Automat. Contr.*, vol. 20, pp. 354–359, 1975.

[33] M. Sternad and A. Ahlén, "The structure and design of realizable decision feedback equalizers for IIR channels with colored noise," *IEEE Trans. Inform. Theory*, vol. 36, pp. 848–858, 1990.

[34] M. Sternad, A. Ahlén, and E. Lindskog "Robust decision feedback equalization," in *Proc. ICASSP'93, Int. Conf. Acoustics, Speech and Signal Processing*, Minneapolis, MN, vol. III, April 26–30, 1993, pp. 555–558.

[35] M. Sternad and A. Ahlén, "A novel derivation methodology for polynomial-LQ controller design," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 116–121, 1993.
[36] ——, "Robust filtering and feedforward control based on probabilistic descriptions of model errors," *Automatica*, vol. 29, pp. 661–679, 1993.
[37] M. Sternad, K. Öhrn and A. Ahlén, "Robust H_2 filtering for structured uncertainty: the performance of probabilistic and minimax schemes," Rep. UPTEC 94090, Systems and Control Group, Uppsala University; submitted to the European Control Conference, Rome, Sept. 1995.
[38] K. S. Vastola and H. V. Poor, "Robust Wiener-Kolmogorov theory," *IEEE Trans. Inform. Theory*, vol. 30, pp. 316–327, 1984.
[39] L. Xie and Y. C. Soh, "Robust Kalman filtering for uncertain systems," *Syst. Contr. Lett.*, vol 22, pp. 123–129, 1994.
[40] L. Xie, Y. C. Soh, and C. E. de Souza, "Robust Kalman filtering for uncertain discrete-time systems" *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1310–1314, June 1994.



Kenth Öhrn (S'92) was born in Västerås, Sweden, in 1960. He received the M.Sc. degree in engineering physics in 1991 from the Institute of Technology, Uppsala University, Sweden. He is wroking toward the Swedish Licentiate degree from the same institution.

He works on control problems at ABB Traction AB, Electrical Division, in Västerås, Sweden. His current research interests include robust deconvolution and robust feedforward control.

Anders Ahlén (S'80–M'89–SM'90) received the Ph.D. degree in automatic control in 1986 from Institute of Technology, Uppsala University, Sweden.

Dr. Ahlén held positions as Assistant and Associate Professor of Automatic Control at Uppsala University from 1984–1989 and 1989–1992, respectively. He spent 1991 with the Department of Electrical and Computer Engineering, University of Newcastle, Australia. From July 1992 he has been Associate Professor of Signal Processing at the Systems and Control Group, Uppsala University. His research interests are currently focused on the robust and adaptive systems in the area of signal processing, communications, and control.

Mikael Sternad (S'83–M'86–SM'90) received the M.S. degree in engineering physics in 1981 and the Ph.D. degree in automatic control in 1987, both from Institute of Technology, Uppsala University, Sweden.

Dr. Sternad is an Associate Professor at the University of Uppsala, Sweden. He has worked with the Systems and Control Group at Uppsala from 1981–1986 as teaching assistant and from 1987as Assistant Professor and Associate Professor. His current research interests include the use of polynomial equations approach on problems in control, estimation, digital communications and the design of adaptive tracking algorithms for systems with fast time-variations.